the number of commodities, or other conditions mentioned at the beginning, and differences due to differences in formulae almost disappear if we get rid of even the crudest bias.

Properly used index numbers are a precise tool, precise far beyond our needs, precise within much less than 1 per cent.

DISCUSSION

By Wesley C. Mitchell

The index number which Professor Fisher proposes seems to me excellent. Indeed, it is "the best index number" known to me for the particular purpose he has in view. Further, that purpose is an important one, covering as it does most of the uses to which what I call "general-purpose" index numbers are commonly put. But I cannot admit that even perfect adaptation of an index number to any use however important entitles it to be called "the best index number" at large.

This issue is much more than a verbal quibble. So far, most makers of index numbers of prices at wholesale have been designing "general-purpose" series. But the time has come when we are beginning to make a wide variety of index numbers especially adapted to particular uses. And we must go further in that direction. In doing such work the compiler should first define as accurately as may be the use which his new series is to serve, and from this use he should deduce the form of index number which is "the best" for him. This criterion of use should determine the number of commodities to be included, the basis on which commodities are selected, the scheme of weighting, and the type of average. What is best in all these respects for one use may be bad for another use. There cannot conceivably be an index number that is "the best" for all uses.

In fine, I think Professor Fisher has tarnished somewhat his excellent contribution by using incautious language implying universal merits in his formula. His very eminence in this progressive field of work makes that slip important. The path of future progress lies in differentiation. Yet if we took Professor Fisher at his word, we should never produce anything but one admirable kind of "general-purpose" series.

DISCUSSION

By C. M. Walsh

In the first place I wish to make an acknowledgment of appreciation for the flattering remarks with which Professor Fisher has referred to my book. In return I must express my wonder at the stupendous labor he has performed in applying so many complex systems of

index numbers to the multitudinous data of the War Industries Board. It has been purely a labor of love, and it is truly admirable. But I could wish he had curtailed it in one direction, and had extended it in another. He might, indeed, have nearly halved his labor without loss. He has shown us that there are four elemental forms of weighting—q_0p_0, q_0p_1, q_1p_0, q_1p_1—two of which are mixtures of the quantities of one period with the prices of another period. These are extremely artificial, and what has happened might have been anticipated—that nothing has been gained by experimenting with them. The other two, which use in their natural connection the quantities and prices of the same period—the one, those of the base period; the other, those of the subsequent period—are the only important forms of weighting; and their employment together constitutes what I have called double weighting. Now, Professor Fisher has well shown that these latter two different weights taken singly—in the case at least of the aggregate system—have opposite defects, the one giving a bias above the truth, and the other a bias below it. This at once suggests taking a mean between them. But the mean between them can be drawn in either of two ways. The one is to carry out each aggregate system by itself, and then to draw a mean between their results. This method is now exclusively used by Professor Fisher. The other is to draw a mean between the two sets of quantities (those of every commodity at the two periods), and then to use the mean quantities in a single aggregate system. This method I almost exclusively used in my book. We both have committed opposite mistakes. It may be—and I am now of this opinion—that the first is the better method. But it needs to be proved that it is. Had Professor Fisher included the other method in his investigations, he might have thrown light on this question.

Consider these two formulae of the aggregate system:

$$(1)\quad \frac{\Sigma q_0p_1}{\Sigma q_0p_0}, \qquad\qquad (2)\quad \frac{\Sigma q_1p_1}{\Sigma q_1p_0}.$$

The first was advocated by Laspeyres in 1871. The second was adopted by Paasche in 1874. Nothing can be offered in proof of the superiority of the one over the other. They are both wrong, straddling the truth on opposite sides, as clearly appeared in some examples I used in my book, and as has now been demonstrated by Professor Fisher. In his *Elements of Statistics*, published in 1901, Dr. Bowley used the arithmetic mean between them, in this form $\frac{1}{2}\left(\frac{\Sigma q_0p_1}{\Sigma q_0p_0}+\frac{\Sigma q_1p_1}{\Sigma q_1p_0}\right)$; but he said merely that this "may be" regarded as a "first approximation," whatever that may mean. Previously, in 1887, the British Association Committee on Index Numbers had advised using the arithmetic mean

between the quantities of the two periods, in this form $\frac{\Sigma\frac{1}{2}(q_0+q_1)p_1}{\Sigma\frac{1}{2}(q_0+q_1)p_0}$ (in which, of course, the two $\frac{1}{2}$'s may be eliminated). But they damned their own recommendation by declaring that, after all, weighting is of little consequence. Hence no attention has been paid to either of those index numbers, although they are much better than others that are in use. It is as important to give a good argument as it is to suggest a good thing. Indeed, the good argument is of first importance, for without it the goodness of the thing will not be perceived. Professor Fisher now holds that the geometric mean should be employed in place of the arithmetic as used by Bowley. The superiority of the geometric mean for this purpose I pointed out in my book; and I am glad that Professor Fisher has come to recognize it. I used it in place of the arithmetic as used by the British Association Committee, in this form $\frac{\Sigma\sqrt{q_0q_1}\cdot p_1}{\Sigma\sqrt{q_0q_1}\cdot p_0}$. This was one of my three best index numbers, which Professor Fisher used among his calculations, and the results of which he threw on the screen. The index number which Professor Fisher now advocates is this: $\sqrt{\frac{\Sigma q_0p_1}{\Sigma q_0p_0}\times\frac{\Sigma q_0p_1}{\Sigma q_1p_0}}$. I had referred to this only in a footnote, where, indeed, it stood the test to which I was subjecting various index numbers better than did my three. Recently I have gone over my work, and last year I wrote a pamphlet on the subject, which is not yet published, but which will probably be out within a month. In it I have taken up this very index number and added it to my other three as a fourth theoretically good one, and perhaps *the* best. But this index number was first formulated by Professor Fisher, and has been first published by him. Therefore, it should be known as Fisher's index number.

Now, how do we know that this is a theoretically good system, and perhaps the best one, or the "ideal" one, as Professor Fisher calls it? Here Professor Fisher's method and mine are very different. But I can now speak only of the tests to which it and others are subjected. In the abstract sent to us Professor Fisher used three tests. I intended to criticize one of them as not a test at all, because a test is a trial of the index number on given examples, whereas this is only a requirement. It is the demand for double weighting which I made in my book, and which Professor Fisher now repeats. In his address as delivered he makes this demand, but does not call it a test. He now has only two tests. The one is a special form of Westergaard's test. Westergaard's test is that a result obtained at a distant period

by going over all the intervening periods, comparing each with its predecessor, and piecing the results together in one sequence, should agree with the result obtained by directly comparing the last with the first. But if the results disagree, Westergaard had no way of knowing which is the better one, much less which is the right one. I therefore altered his test by adding as the last period one exactly like the first, both in quantities and in prices. Then we *know* that the price-level and the exchange-value of money are the same as at the first period, and this sameness, which will always be indicated by the direct comparison, ought to be indicated also by the index number employed in comparisons among all the intervening periods, in a string or chain spanning the interval. It happens that even bad systems, which fall down before other simpler tests, will yield the right result at the end of such a chain, while they yield very differing results for the intervening periods. This test, like all other tests, in fact, is only a negative test. It cannot prove an index number, but it can disprove an index number.

It happens also that our very best systems, which meet most of the other tests and agree most closely with one another, only come near to giving the known result at the end of such a chain; but evidently the one that comes nearest, always or nearly always, is the least erroneous. Herein consists the value of this test. It is like the test to which surveyors submit their measurements when they have gone round a field and, coming back to their starting point, require that it should be accurately indicated. It may, therefore, be called the circular test. But in its use an important distinction must be observed. For its full efficacy this test must be applied to at least four periods, at least two intervening periods being different; for then we have at least two transitions that are different, and we come back to our starting point over a different track from that over which we went out; whereas, if it is applied to only three periods, with only two transitions, these transitions are exactly alike except for the mere fact of being reversed. It is true that some index numbers will not stand the test even confined to three periods; but there are others that will stand it so confined, but will fail when it is extended to more than three periods; wherefore this is a more accurate use of the test. This peculiarity attaches to other tests as well, and it was recognized by Professor Fisher when he wrote his book; for in his book he gave a half-mark to the index numbers that satisfy his tests only over three periods and a full mark only to those which satisfy them over more than three periods. Now here he has used this additional circular test—Westergaard's test as altered by me; and he has used it the first time it has ever been

used since I made the alteration twenty years ago. But he has applied it to only three periods. You may be surprised at this assertion, as you can recall no reference or allusion made by Professor Fisher to three periods. But in going from a first period to a second, and then back to the first, and demanding agreement, Professor Fisher has done the same as going on from the second to a third period exactly like the first. But why has he not extended this test to more than three periods? The "ideal" system which he recommends fulfils this test confined to three periods, but does not fulfil it extended to more than three periods. I should be the last person to think that this was a reason for ignoring the full test. Why he has made the confinement, I do not understand.

As for the last test, which Professor Fisher has numbered the first, it is an entirely new one, original, and highly ingenious. I refer to the requirement that when an index number, applied first to the prices, is then turned about and by interchanging the price and quantity symbols is applied to the quantities, the two results multiplied together should yield the total value variation indicated by $\frac{\Sigma q_1 p_1}{\Sigma q_0 p_0}$. This test seems to have some merits, and perhaps might be employed to measure the relative accuracy of theoretically good index numbers. But I have not time to inquire into its merits. Professor Fisher now seems to think it a very important test. In his abstract he seemed to use it almost as the crucial test. He there wrote that as far as he knew the "ideal" index number which he recommends is the only one that fulfils this test. That sounded like an argument for this index number; and I wrote a criticism demolishing it. But yesterday Professor Fisher told me he was aware that there are other index numbers that fulfil this test. What I wrote, therefore, can no longer serve as a criticism. Still, Professor Fisher may not be aware of the great number of index numbers that fulfil this test—a fact which considerably diminishes its importance; and therefore I will read what I wrote, which presents him with at least twenty such index numbers. Not that I intend to describe all these index numbers to you. I will show you only the method of obtaining them.

The "ideal" index number, now recommended at least for its not failing to meet this new test, was formulated, as I have said, in Professor Fisher's book, and it appears in his list with the double number 15 and 16. It is doubly numbered because Professor Fisher points out that it has "the distinction of being identical with its own antithesis." If Professor Fisher will examine more closely into the matter, I believe he will find that for an index number to be identical with its

own antithesis is the same as for it to fulfil this new test. This distinction did not appeal to him then as it does now. But, unless I misunderstand his correlation of price and quantity indexes, this distinction is only apparent, owing to our not having all the formulae of index numbers that we might have. I take it that most of the persons present know that Professor Fisher's list of index numbers in *The Purchasing Power of Money* contains forty-four, arranged in twenty-two pairs of antithetical formulae. The import of this arrangement I have never been able to fathom, especially since, as I have privately pointed out to Professor Fisher, some of the antithetic averages—some of the ones with even numbers—are not averages at all; and his proof that they must be rests on a wrong definition of average. Yet it is interesting to learn that some of the old familiar formulae admit of being thus opposed to each other. Now, the system in question, constituting a pair by itself, numbered 15 and 16, is the geometric mean between the two systems numbered 11 and 12, which together form one pair. It is the only formula in the list that thus draws the geometric mean between the two members of any pair; and this, I think, will turn out to be the reason why it has the distinction referred to. We are not restricted to drawing the geometric mean between index numbers 11 and 12: we can draw it between the members of any other pair, and set the resultant formula up as an index number. I note that the formulae numbered 26 and 28 each contains a misprint, which must be corrected before the right result can be obtained. I throw out the last pair, Nos. 43 and 44, because these are really not averages at all. I also reject the pair consisting of Nos. 9 and 10, because these contain the median, which I for one do not recognize as a mean. The pairs Nos. 11 and 12 and Nos. 15 and 16 having already been accounted for, there remain eighteen pairs. Now, if Professor Fisher will draw the geometric mean between the index numbers in all these eighteen pairs, I believe he will find that in every case the resultant formula will be identical with its own antithesis and will fulfil the new test. Indeed, I believe it to be demonstrable that always the geometric mean between antithetical index numbers must have these properties. And even this is not all; for if he will go on and draw the geometric mean between the formulae obtained from the second and third pairs, and again between those obtained from the thirteenth and seventeenth, he will, I believe, obtain two more index numbers possessing the same properties. We thus have a method for obtaining all the index numbers we please that will be their own antitheses and will fulfil the new test. We merely have to take any index number, find its antithesis in the way prescribed by Professor Fisher, and then draw the geometric mean between the two.

However, I agree with Professor Fisher in considering the index number he calls the "ideal" one the best one. And now, in closing, I will make an appeal for harmony; and I will make an appeal to Professor Fisher against himself. For if I have rightly grasped his conclusion, which was somewhat hurriedly read, he has ended, in spite of his recommendation of this very workable "ideal" index number, by telling us that for practical purposes either Laspeyres' or Paasche's index number is good enough, and that, therefore, as Laspeyres' index number, which uses only the quantities of the base year once for all, is the easier, it should be employed. In his book he recommended Paasche's index number, now he recommends Laspeyres'—its counterpart! If this is the result of all his labors, I regard it as very regrettable indeed. It is true that the figures yielded by these two index numbers applied to the data supplied by the War Industries Board show very slight divergencies from each other and from those yielded by the index numbers which draw a mean between them. But we should note that his investigation has continued over only four years, whereas he recommends these systems for an indefinite term of years. Yet it is a peculiarity of these two index numbers, as of several other bad ones, that their errors accumulate, so that in a long sequence they may become quite appreciable. But in any case, why employ an inferior index number when we possess a better? Mere convenience is not a good reason, and after this display of his patience and perseverance, Professor Fisher is the last one from whom we should expect a compromise with sloth. Perhaps the worst feature in the matter is that until all adopt the best method there never will be agreement; for if you once admit a compromise for mere convenience, there will be different compromises, and the confusion which now reigns in our subject, which is a disgrace to statistical science, will continue. You, gentlemen of the American Statistical Association, have a duty to perform with regard to this subject. Many of you are engaged in statistical work. Many of you advise others who are or will be so engaged. Adopt and recommend the best. Now, in researches into the past the data for a yearly comparison of quantities may not be forthcoming. Then, and then only, you may compromise because you are compelled to do so by the circumstances in the case. You can do only the best the circumstances allow. What index number should you then use? It should be this: $\frac{\Sigma q p_1}{\Sigma q p_0}$. This is the method used by Lowe within a year or two of one hundred years ago. In my book I called it Scrope's index number; but it should be called Lowe's. Note that in it are used quantities neither of a base year nor of a subsequent

year. The quantities used should be rough estimates of what the quantities were throughout the period or epoch. For instance, if you are trying to get the course of the price-level since 1800, you might begin with rough estimates of the quantities for the period 1800–1815 —a war period. After that, another set of estimates of the quantities should be used, which might last for ten years. Ten-year terms might be used down to 1875, after which five-year changes might be introduced. These should be still further shortened as you come nearer to the present. For recent years—and especially for the continued use of index numbers into the future—the quantities of every year should be used, and a mean should somehow be drawn between those of every two years. If the geometric mean is antipathetic to you, use the arithmetic, in Bowley's system. The difference between it and the geometric will be but trifling. But, after all, the geometric mean, being used only once in each yearly computation, is by no means beyond the amount of industry we have a right to expect from those who labor with figures. I have no quarrel with Professor Mitchell because, in editing the immense work on the *History of Prices During the War*, he attempted to use Paasche's index number and succeeded in using it only for the first three years, falling into the use of Laspeyres' for the last year, all on the quantities of an intermediate year. He was in a great hurry, and had a right to choose a simple system. But if his work is to become a permanent governmental function—if perchance it should ever be made the guiding principle for Professor Fisher's own scheme of stabilizing the dollar—then we want the very best possible system. Remember then: for the past, use Lowe's index number, and for the present and the future, use Fisher's.

DISCUSSION

By Warren M. Persons

Professor Fisher's analysis of index numbers enables us to use them with greater confidence than we formerly had in the reality of the results. He has shown that they are instruments of precision.

You may have noticed that Professor Fisher has discussed the method of averaging and of weighting index numbers of prices and production without referring explicitly to the *purpose* for which those index numbers were constructed. It seems to me that the methods of averaging and of weighting, as well as the selection of the basic data, all depend upon the purpose to which the index number is to be applied. Moreover, it is evident that Professor Fisher had a specific purpose in mind. His "ideal" index number is an index number that fits into the equation of exchange. He holds that the product of a